Measuring the cardiac output in acute emergency admissions: use of the non-invasive ultrasonic cardiac output monitor (USCOM) with determination of the learning curve and inter-rater reliability

Luke E Hodgson[1,2], Richard Venn[1], Lui G Forni[3], Theophilus L Samuels[3] and Howard G Wakeling[1]

Abstract

Traditionally, assessment of the cardiac output has been limited to theatre or the intensive care unit. However, non-invasive cardiac output estimation is now readily available, and its application may have wider benefit in the emergency setting. The non-invasive ultrasonic cardiac output monitor (USCOM) was investigated to determine its learning curve and inter-rater reliability. Four trainee operators each performed stroke volume measurements on 25 volunteers, compared to an experienced operator pre- and post-passive leg raise. Inter-rater reliability was then assessed on 24 acute emergency in-patients. Mean percentage difference in stroke volume decreased from 19% (95% confidence intervals 14–23) across volunteers 1–5, to 6% (4–8) for the last 5 volunteers scanned. Consequently, on acute emergency in-patients, excellent inter-rater reliability (Lin's concordance correlation coefficient (ρ_c) 0.96 (0.92–0.98)) and agreement of a change $\geqslant 10\%$ in stroke volume following passive leg raise on 23/24 cases were found. Following a training period of less than 5 h, USCOM stroke volume measurements demonstrated excellent inter-rater reliability.

Keywords

Measurement techniques, cardiac output, equipment, ultrasound machines

Introduction

Basic haemodynamic parameters (such as blood pressure), examination, imaging (chest X-ray), functional measures of organ perfusion (such as urine output) and the central venous pressure are routinely used to guide therapy, but do not accurately predict volume status or responsiveness.[1–3] Given that injudicious fluid usage leading to volume overload is associated with increased morbidity and mortality,[4,5] measures that may prevent this should be explored. Titrating fluid resuscitation guided by stroke volume (SV) changes reflects physiological response.[1,3,6] Using such a goal-directed approach has been shown to reduce morbidity and hospital length of stay in patients undergoing surgery,[7] though recent trials have raised questions around the approach.[8–11] Non-invasive cardiac output (CO) devices are now available that may have wider utility. For instance, in the first few hours of hospital attendance, when managing haemodynamic instability may

be of greatest importance, or on a ward without access to invasive devices.

The Doppler equation has been applied to measure blood flow for over half a century. The Ultrasonic Cardiac Output Monitor (USCOM 1A, Uscom Ltd., Sydney, NSW, Australia) is a non-invasive continuous wave (CW) Doppler ultrasound device. CW Doppler

[1]Anaesthetics & Intensive Care Department, Western Sussex Hospitals NHS Foundation Trust, Worthing Hospital, Worthing, UK
[2]Primary Care and Population Sciences, University of Southampton, Southampton General Hospital, Southampton, Hampshire, UK
[3]Intensive Care Department, The Royal Surrey County Hospital NHS Foundation Trust, Guildford, Surrey, UK

Corresponding author:
Luke Eliot Hodgson, Intensive Care Department, Western Sussex Hospitals NHS Foundation Trust, Worthing Hospital, Lyndhurst Rd, Worthing, Western Sussex BN11 2DH, UK.
Email: drlhodgson@gmail.com

has an accuracy measured by Doppler string phantoms of $\pm 2.3\%$.[12] Placed in the supra-sternal notch, the USCOM measures Doppler flow at the aortic valve and determines this over a cardiac cycle (the velocity time integral (VTI)). SV is calculated by multiplying the VTI by an estimate of valve cross-sectional area, determined by applying height-indexed regression equations (if the valve area is known this can be manually imputed).[13] USCOM has been compared positively with the gold standard measurement of CO, the aortic flow probe, in animals[14,15] and in externally driven artificial hearts in orthotopic transplantation with the pulmonary artery catheter.[16] For such a device to be useful clinically, it must be accurate, precise, detect significant directional changes and be reliable.[17]

A passive leg raise (PLR) is a simple, reversible manoeuvre that mimics a rapid volume expansion by shifting blood from the extremities with maximal effects within 90 s. A meta-analysis demonstrated that PLR-induced change in SV represents a more sensitive, specific predictor of fluid responsiveness than changes in arterial pulse pressure on ventilated patients.[18] There have been few studies of PLR on non-intubated patients, though Preau et al.[19] on 34 patients with sepsis or pancreatitis, found a change (Δ) in SV $\geqslant 10\%$ induced by a PLR accurately predicted a significant increase in SV following volume expansion (defined as a SV of $\geqslant 15\%$), with an area under the receiver operating curve of 0.94. Thiel and colleagues[20] demonstrated the utility of the USCOM in conjunction with a PLR to predict fluid responsiveness in critically ill patients. None of the published inter-rater reliability studies of this device have included a PLR (or any intervention), and the majority of studies have been on paediatric populations.[21–27] Dey and Sprivulis[21] proposed the Freemantle criteria (Table 1) to assess acoustic image quality using the USCOM during training, to ensure a reliable image is selected, with a learning curve of 20 readings reported. The other inter-rater studies did not focus on learning curve skill acquisition.

The two objectives of this study were to document the learning curve for the USCOM device for novice operators (doctors in training), assessed against an experienced operator, using volunteers (including a PLR); secondly to assess inter-rater reliability, whilst performing a PLR, on acute hospital patients.

Table 1. Freemantle criteria (out of six) to assess image quality using the USCOM device from Dey & Sprivulis.[22]

1. Well-defined image base
2. Well-defined image peak
3. Well-defined commencement of flow or heart sound
4. Well-defined cessation of flow or heart sound
5. Appropriate scale used on screen
6. Minimal acoustic interference

Methods

This observational single-centre study was conducted at a District General Hospital in the UK. A favourable
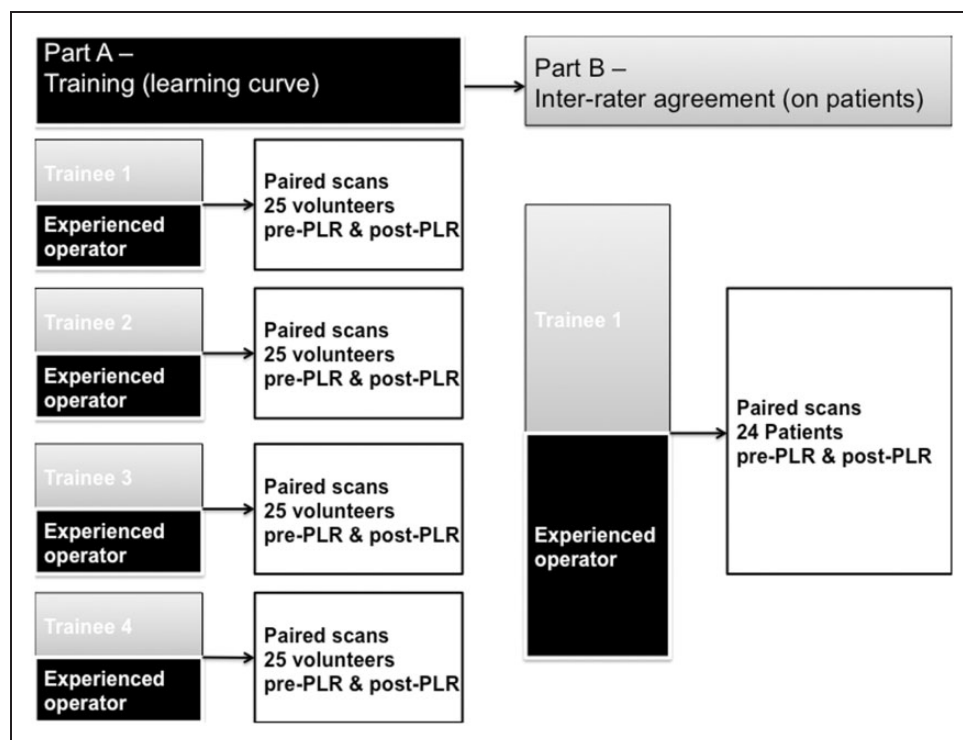

Figure 1. Study flow chart. Part A: learning curve assessed with $n = 4$ trainees, $n = 25$ volunteers. Part B: inter-rater reliability on 24 in-patients.
PLR: Passive leg raise.

opinion was received from the North West Preston Ethics Committee (REC reference 13/NW/0437). Informed written consent was obtained. Patients were excluded if under 18 years, or if they had a history of valvular disease or significant dysrhythmias. The two-part nature of the study is outlined in Figure 1.

Learning curve

The learning curve was assessed using four trainee operators (4–6 years post-medical qualification), who were initially given a 30-min lecture of how to use the device, including a review of the basic physics of ultrasound and the CW Doppler technique. A priori, for clinical usefulness, an inter-rater difference of 5% was felt to represent acceptably close agreement, taking into account normal beat-to-beat changes in SV. Each trainee gained SV values pre- and post-PLR on healthy volunteers ($n = 100$, mean age 40, range 18–65), in a stepwise, blinded fashion, alongside an experienced operator (>200 scans performed). Sample size was based on the previous documentation of a learning curve of 20 scans[21] – as the study also involved a PLR, it was decided to enlist 25 volunteers for each novice operator, to provide in total 50 scans (25 pre- and post-PLR). Each reading was performed with the full focusing technique with the Doppler flow profile held steady for a 7.5 second sweep. Aberrant flow profile outlines on visual inspection were rejected, and scans scored according to the Freemantle criteria (Table 1).[21]

In an iterative process, the trainee was given feedback on image optimisation following each pair of readings. In all, 100 healthy volunteers were scanned, providing 200 paired readings. Different volunteers were used to train each trainee due to resource availability. For the PLR, a subject is placed semi-recumbent for 3 min, then supine with the legs elevated to 45°. Total training time including data collection pre- and post-PLR on 25 volunteers by the two personnel, took 4–5 h.

Inter-rater reliability

This part of the study assessed inter-rater agreement for absolute SV as well as detection of a $\geqslant 10\%$ change (Δ) in SV induced by PLR. There is no consensus on what constitutes a significant Δ following PLR, or Δ in SV following volume expansion to conclude a patient is fluid 'responsive'. However, most studies and expert recommendations use a cut-off between 10% and 15%.[6,28–30] For example, in a study by Monnet et al.,[31] a $\Delta 10\%$ induced by PLR best predicted their definition of a significant Δ in SV – in this case, a $\Delta 15\%$ in SV following volume expansion. Pre- and post-PLR scans were performed, in a blinded fashion, on 24 in-patients (mean age 55, range 22–88; 15 male, 9 female) on acute medical and surgical wards, by one newly trained operator and the experienced operator. Both operators were blinded to the patients' diagnosis and history. As in a previous study[22] (with a similar sample size), only one of the four trainees participated, as it was felt that including more would be too onerous for patients, and there was limited availability of trainees to take part concurrently.

Statistics

A wide range of statistical techniques have been described to assess the learning curve in clinical situations, though none have been described in reference to the USCOM.[32,33] Mean percentage difference in SV measurements obtained by the paired trainee, compared to the experienced operator pre- and post-PLR, was calculated on each volunteer scanned. Following this, a linear model with 95% confidence intervals was fitted to the data, in an attempt to visually demonstrate convergence in the mean percentage differences around the fitted line as the number of scans (and experience) increased.

The experienced operator was taken to represent the gold standard for SV values using the device, over the course of the scans. The concordance correlation coefficient (ρ_c) proposed by Lin[34] for assessment of concordance in continuous data avoids the shortcomings associated with a number of alternative procedures, including Pearson correlation coefficient r, paired *t*-tests, coefficient of variation and intra-class correlation coefficient. It is said to be robust on as few as 10 pairs of data. This technique was used to compare each trainee, with the experienced operator, during scans on the first five and last five volunteers and subsequently to assess inter-rater reliability on patients.[34] A value of one denotes perfect concordance; a value of zero denotes complete absence. Ease of assessment and the Fremantle score were also recorded.

All data were imported into an Excel^{TM} spreadsheet (Microsoft, Redmond, WA, USA) used to calculate mean differences. All other statistical analyses were performed using the open source statistical package R (GNU project, R Foundation for Statistical Computing, Vienna, Austria).

Results
Learning curve

During training, across 25 volunteers, mean percentage difference in SV readings ($n = 200$) between novice trainees ($n = 4$) and experienced operator decreased significantly from 19% (95% confidence intervals 14–23) during scans 1–5, to 5.9% (4–8) for the last five volunteers (Figure 2). Lin's concordance correlation coefficient (ρ_c) increased from 0.75 (0.62 to 0.84) for the first five volunteers to 0.91 (0.85 to 0.95) for the scans from volunteers 21–25.
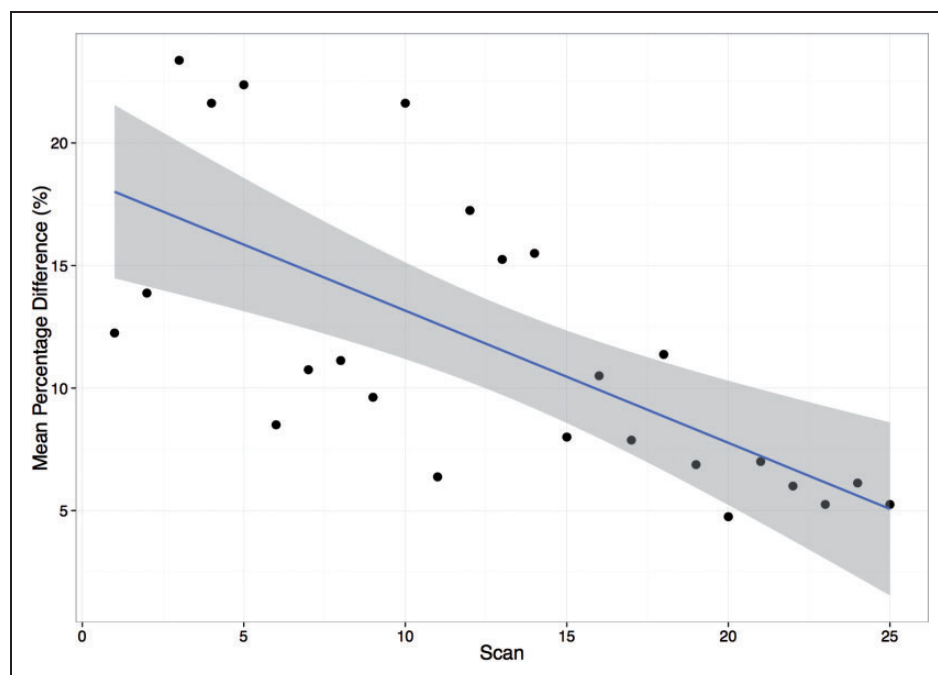
Figure 2. Learning curve on 25 volunteers (50 scans – 25 pre- PLR and 25 post-PLR). Mean percentage difference SV values between trainees ($n = 4$) and experienced operator decreased significantly from 19% (95% confidence intervals 14–23) during scans 1–5, to 5.9% (4–8) for the last five volunteers during training. The shaded area represents the 95% confidence interval. PLR = passive leg raise. SV = stroke volume.

Inter-rater reliability

Inter-rater reliability was assessed on 24 patients pre- and post-PLR. Mean SV pre-PLR was 61 ml (54–67) for the experienced operator and 59 ml (53–66) for the newly trained operator; mean percentage difference in SV between operators was 5.8%. Post-PLR SV values were 76 ml (68–83) and 74 ml (66–82), respectively, with a mean percentage difference of 4.8%. Across all of the patient scans, Lin's concordance correlation coefficient (ρ_c) was 0.96 (0.92–0.98) (Figure 3). On 23/24 occasions, the two operators agreed as to whether a PLR had produced a $\geqslant 10\%$ increase in SV. Four of the patients did not significantly increase SV with a PLR, all with a history of cardiac disease. Freemantle score for image quality was $\geqslant 4/6$ for all 24 patients, with 83% scoring $\geqslant 5/6$.

Discussion

Following a training period of 50 scans on 25 healthy volunteers, inter-rater reliability showed excellent correlation using the non-invasive USCOM on 24 acute in-patients, using a PLR. This study adds to the existing literature on inter-rater studies of the device (summarised in Table 2). This is the first learning curve and inter-rater study to include a PLR manoeuvre, an intervention that mimics a fluid challenge to the central circulation. Dey and Sprivulis[21] described training that included image quality acquisition and use of both aortic and pulmonary windows. Nguyen et al. assessed inter-rater agreement without a training period.[22] In paediatric[24,25,27] and neonates[26] good inter-rater reliability has been described. Indeed, in a large study of Chinese paediatric patients, Cattermole et al. found the USCOM to have less inter-rater variation than standard automated oscillometric device for blood pressure or heart rate.[24]

The first step in the haemodynamic management of acutely unwell patients is to determine adequacy of tissue perfusion. At present, assessment of response to therapy outside theatre and critical care relies on soft endpoints such as examination, arterial pressure or urine output,[1–3] potentially leading to inadequate resuscitation, or overload. SV directly impacts on oxygen delivery, and it follows that optimising SV should play a role in the management of acutely unwell patients. The work presented shows that the non-invasive USCOM has a relatively short learning curve and could be studied in a clinical trial as part of a novel algorithm applied outside the critical care environment, for example, in the early management of sepsis. Such trials could be argued for given the relative dearth of outcome studies using non-invasive devices and recent negative trials for early goal-directed therapy in sepsis.[35]

The PLR accurately predicts SV response to a fluid challenge,[18] thus potentially avoiding unnecessary fluid loading. As maximal haemodynamic effects of the PLR occur rapidly, it is important to assess these effects with a device on a real-time basis.[31] This study demonstrated that trainees could rapidly attain a SV pre- and post-PLR, adding to the work of Thiel and colleagues.[20]
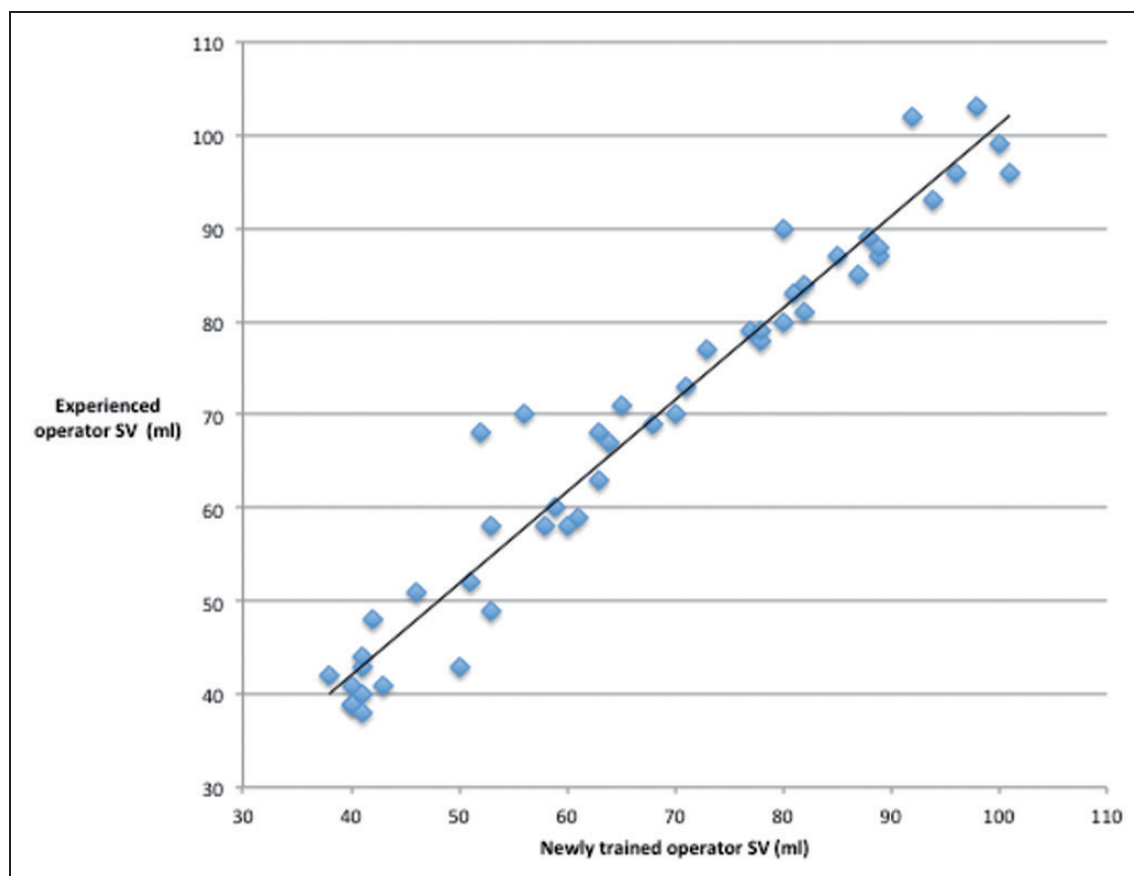
Figure 3. Inter-rater agreement – absolute SV values pre- and post-PLR on 24 patients, between a newly trained and experienced operator. Lin's concordance correlation coefficient (ρ_c) 0.96 (95% confidence intervals 0.92–0.98). PLR = passive leg raise. SV = stroke volume.

Table 2. Existing inter-rater studies.

Study	Population	Training	Raters (sample size)	Correlation statistics
Dey (2005)	ED Adults	20 volunteers	2 ($n=21$)	$r=0.96$ (0.90–0.98)*
Nguyen (2006)	ED Adults	Brief overview	204 ($n=91$)	$r^2=0.87$ (0.86–1.00)
Stewart (2008)	ED Paediatrics	20 volunteers	5 ($n=97$)	$r=0.79$ (0.70–0.86)*
Meyer (2009)	Neonates	20 volunteers	2 ($n=12$)	$r=0.93$ (0.86–0.97)
Cattermole (2010)	Paediatrics	–	2 ($n=1059$)	$r=0.94$ (0.93–0.94)
Dhanani (2011)	Paediatrics (theatre)	20 volunteers	2 ($n=59$)	0.87 (0.79–0.92)**
Napoli (2012)	ED Adults	20 volunteers	2 ($n=30$)	$r=0.95$

ED = emergency department, (95% Confidence intervals). *Pearson product moment correlation coefficient; **Lin's concordance correlation coefficient.

Limitations

Focusing of the probe (visual and acoustic) to obtain optimal and maximal Doppler signal plays a critical role in all Doppler devices[36] and, although the learning curve has been quoted as twenty scans for the USCOM device,[21] in our experience, it takes longer to become familiar with the nuances of different signal sounds and patterns to recognise a reliable signal. Improper technique and poor beam alignment with flow at the aortic outflow tract leads to suboptimal VTI measurements and underestimation of SV.

Our study has weaknesses. The learning curve was performed on healthy volunteers, younger than the average hospitalised patient. Due to resource and time constraints, a different set of volunteers was used for each trainee and only one trainee was subsequently assessed against the experienced operator, on acutely admitted in-patients. As the study involved a PLR, where maximal effects are reached and then wane rapidly, it was not feasible to assess more than one trainee at a time on the volunteers and patients studied. As previously noted, elderly patients can be more difficult to obtain good quality images from.[37,38] However, the

inter-rater study included a number of elderly patients with high correlation between raters. The study did not assess whether acquisition of the skill was sustained, however, as with most simple practical skills, in our experience, this has not been problematic. No interventions were performed on patients, however, the PLR has been shown to be an accurate reflection of a fluid challenge in a number of studies, and this is the first study to assess inter-rater agreement of the device with such a manoeuvre.[18,31]

Conclusions

The learning curve for the USCOM device shows proficiency is gained in fewer than 50 scans; this can be achieved in less than 5 h, with close supervision. Following training, the device shows excellent inter-rater reliability. Adoption of such techniques could improve assessment of the circulation in our unstable patients. Future studies to assess the impact of using such technology, as part of a goal-directed approach, on patient outcomes are desirable.

Declaration of Conflicting Interests

Funding

References

1. Saugel B, Ringmaier S, Holzapfel K, et al. Physical examination, central venous pressure, and chest radiography for the prediction of transpulmonary thermodilution-derived hemodynamic parameters in critically ill patients: a prospective trial. *J Crit Care* 2011; 26: 402–410.
2. McGee S, Abernethy WB, and Simel DL. The rational clinical examination Is this patient hypovolemic? *JAMA* 1999; 281: 1022–1029.
3. Marik PE, and Cavallazzi R. Does the central venous pressure predict fluid responsiveness? An updated meta-analysis and a plea for some common sense. *Crit Care Med* 2013; 41: 1774–1781.
4. Vincent JL, Sakr Y, Sprung CL, et al. Sepsis in European intensive care units: results of the SOAP study. *Crit Care Med* 2006; 34: 344–353.
5. Wiedemann HP, Wheeler AP, Bernard GR, et al. Comparison of two fluid-management strategies in acute lung injury. *N Engl J Med* 2006; 354: 2564–2575.
6. Michard F, and Teboul JL. Predicting fluid responsiveness in ICU patients: a critical analysis of the evidence. *Chest* 2002; 121: 2000–2008.
7. Grocott MP, Dushianthan A, Hamilton MA, et al. Perioperative increase in global blood flow to explicit defined goals and outcomes after surgery: a Cochrane Systematic Review. *Br J Anaesth* 2013; 111: 535–548.
8. Srinivasa S, Taylor MH, Singh PP, et al. Randomized clinical trial of goal-directed fluid therapy within an enhanced recovery protocol for elective colectomy. *Br J Surg* 2013; 100: 66–74.
9. Brandstrup B, Svendsen PE, Rasmussen M, et al. Which goal for fluid therapy during colorectal surgery is followed by the best outcome: near-maximal stroke volume or zero fluid balance? *Br J Anaesth* 2012; 109: 191–199.
10. Pearse RM, Harrison DA, MacDonald N, et al. Effect of a perioperative, cardiac output-guided hemodynamic therapy algorithm on outcomes following major gastrointestinal surgery: a randomized clinical trial and systematic review. *JAMA* 2014; 311: 2181–2190.
11. Phan TD, D'Souza B, Rattray MJ, et al. A randomised controlled trial of fluid restriction compared to oesophageal Doppler-guided goal-directed fluid therapy in elective major colorectal surgery within an Enhanced Recovery After Surgery program. *Anaesth Intensive Care* 2014; 42: 752–760.
12. Walker A, Olsson E, Wranne B, et al. Accuracy of spectral Doppler flow and tissue velocity measurements in ultrasound systems. *Ultrasound Med Biol* 2004; 30: 127–132.
13. Nidorf SM, Picard MH, Triulzi MO, et al. New perspectives in the assessment of cardiac chamber dimensions during development and adulthood. *J Am Coll Cardiol* 1992; 19: 983–988.
14. Critchley LA, Peng ZY, Fok BS, et al. Testing the reliability of a new ultrasonic cardiac output monitor, the USCOM, by using aortic flowprobes in anesthetized dogs. *Anesth Analg* 2005; 100: 748–753.
15. Phillips RA, Hood SG, Jacobson BM, et al. Pulmonary artery catheter (PAC) accuracy and efficacy compared with flow probe and transcutaneous Doppler (USCOM): an ovine cardiac output validation. *Crit Care Res Pract* 2012; 2012: 621496.
16. Phillips R, Lichtenthal P, Sloniger J, et al. Noninvasive cardiac output measurement in heart failure subjects on circulatory support. *Anesth Analg* 2009; 108: 881–886.
17. Squara P, Cecconi M, Rhodes A, et al. Tracking changes in cardiac output: methodological considerations for the validation of monitoring devices. *Intensive Care Med* 2009; 35: 1801–1808.
18. Cavallaro F, Sandroni C, Marano C, et al. Diagnostic accuracy of passive leg raising for prediction of fluid responsiveness in adults: systematic review and meta-analysis of clinical studies. *Intensive Care Med* 2010; 36: 1475–1483.
19. Preau S, Saulnier F, Dewavrin F, et al. Passive leg raising is predictive of fluid responsiveness in spontaneously breathing patients with severe sepsis or acute pancreatitis. *Crit Care Med* 2010; 38: 819–825.
20. Thiel SW, Kollef MH, and Isakow W. Non-invasive stroke volume measurement and passive leg raising predict volume responsiveness in medical ICU patients: an observational cohort study. *Crit Care* 2009; 13: R111.
21. Dey I, and Sprivulis P. Emergency physicians can reliably assess emergency department patient cardiac output using the USCOM continuous wave Doppler cardiac output monitor. *Emerg Med Australas* 2005; 17: 193–199.
22. Nguyen HB, Losey T, Rasmussen J, et al. Interrater reliability of cardiac output measurements by transcutaneous Doppler ultrasound: implications for

noninvasive hemodynamic monitoring in the ED. *Am J Emerg Med* 2006; 24: 828–835.

23. Napoli AM, Machan JT, Corl K, et al. Interrater reliability and parallel forms reliability of two noninvasive hemodynamic monitoring devices in clinically stable emergency department patients. *Eur J Emerg Med* 2012; 19: 297–303.

24. Cattermole GN, Leung PY, Mak PS, et al. The normal ranges of cardiovascular parameters in children measured using the Ultrasonic Cardiac Output Monitor. *Crit Care Med* 2010; 38: 1875–1881.

25. Stewart GM, Nguyen HB, Kim TY, et al. Inter-rater reliability for noninvasive measurement of cardiac function in children. *Pediatr Emerg Care* 2008; 24: 433–437.

26. Meyer S, Todd D, and Shadboldt B. Assessment of portable continuous wave Doppler ultrasound (ultrasonic cardiac output monitor) for cardiac output measurements in neonates. *J Paediatr Child Health* 2009; 45: 464–468.

27. Dhanani S, Barrowman NJ, Ward RE, et al. Intra- and inter-observer reliability using a noninvasive ultrasound cardiac output monitor in healthy anesthetized children. *Paediatr Anaesth* 2011; 21: 858–864.

28. Bendjelid K, and Romand JA. Fluid responsiveness in mechanically ventilated patients: a review of indices used in intensive care. *Intensive Care Med* 2003; 29: 352–360.

29. Marik PE, Levitov A, Young A, et al. The use of bioreactance and carotid Doppler to determine volume responsiveness and blood flow redistribution following passive leg raising in hemodynamically unstable patients. *Chest* 2013; 143: 364–370.

30. Cecconi M, Parsons AK, and Rhodes A. What is a fluid challenge? *Curr Opin Crit Care* 2011; 17: 290–295.

31. Monnet X, Rienzo M, Osman D, et al. Passive leg raising predicts fluid responsiveness in the critically ill. *Crit Care Med* 2006; 34: 1402–1407.

32. Khan N, Abboudi H, Khan MS, et al. Measuring the surgical 'learning curve': methods, variables and competency. *BJU Int* 2014; 113: 504–508.

33. Ramsay CR, Grant AM, Wallace SA, et al. Statistical assessment of the learning curves of health technologies. *Health Technol Assess* 2001; 5: 1–79.

34. Lin LI. A concordance correlation coefficient to evaluate reproducibility*Biometrics* 1989; 45: 255–268.

35. Angus DC, Barnato AE, Bell D, et al. A systematic review and meta-analysis of early goal-directed therapy for septic shock: the ARISE, ProCESS and ProMISe Investigators. *Intensive Care Med* 2015; 41: 1549–1560.

36. Anderson B. Basic principles of spectral Doppler. In B Anderson (ed.) *Echocardiography: The Normal Examination and Echocardiographic Measurements*. Queensland: MGA Graphics, 2000, pp. 55–66.

37. Huang L, and Critchley LA. Study to determine the repeatability of supra-sternal Doppler (ultrasound cardiac output monitor) during general anaesthesia: effects of scan quality, flow volume, and increasing age. *Br J Anaesth* 2013; 111: 907–915.

38. Huang L, Critchley LA, Lok RL, et al. Correlation between supra-sternal Doppler cardiac output (USCOM) measurements and chest radiological features. *Anaesthesia* 2013; 68: 1156–1160.